

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2004-078929

(43)Date of publication of application : 11.03.2004

(51)Int.Cl. G06F 15/177  
G06F 1/04

(21)Application number : 2003-276005 (71)Applicant : HEWLETT-PACKARD DEVELOPMENT CO LP

(22)Date of filing : 17.07.2003 (72)Inventor : ESPINOZA-IBARRA RICARDO  
BARR ANDREW H

(30)Priority

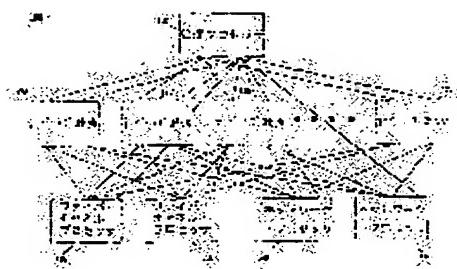
Priority number : 2002 216283 Priority date : 12.08.2002 Priority country : US

## (54) SYSTEM WHICH OPTIMIZES PERFORMANCE IN PROCESSOR LEVEL, METHOD, AND APPARATUS

### (57)Abstract:

PROBLEM TO BE SOLVED: To provide an optimum processor frequency management system in a multiprocessor system.

SOLUTION: In one embodiment of frequency management of a processor in a multiprocessor system, a first processor requiring a first level performance, operates on a special frequency and consumes a part of heat/electric power budget of this chassis. A second processor requiring a second level performance, operates on a special frequency and consumes a part of heat/electric power budget of this chassis. A integrated electric power/heat budget in a multiprocessor computer system is maintained under a different operation frequency.



### LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision  
of rejection]

[Date of requesting appeal against examiner's  
decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁(JP)

## (12) 公開特許公報(A)

(11) 特許出願公開番号

特開2004-78929

(P2004-78929A)

(43) 公開日 平成16年3月11日(2004.3.11)

(51) Int.CI.<sup>7</sup>

G06F 15/177

G06F 1/04

F 1

G06F 15/177 674Z

G06F 1/04 301C

テーマコード(参考)

5B045

5B079

審査請求 未請求 請求項の数 10 O L (全 12 頁)

(21) 出願番号

特願2003-276005(P2003-276005)

(22) 出願日

平成15年7月17日(2003.7.17)

(31) 優先権主張番号

10/216,283

(32) 優先日

平成14年8月12日(2002.8.12)

(33) 優先権主張国

米国(US)

(71) 出願人 503003854

ヒューレット・パッカード デベロップメント カンパニー エル.ビー.  
アメリカ合衆国 テキサス州 77070  
ヒューストン 20555 ステイト  
ハイウェイ 249

(74) 代理人 100081721

弁理士 岡田 次生

(74) 代理人 100105393

弁理士 伏見 直哉

(74) 代理人 100111969

弁理士 平野 ゆかり

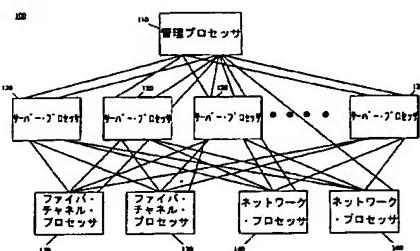
(54) 【発明の名称】プロセッサ・レベルにて性能を最適化するシステム、方法、装置

## (57) 【要約】

【課題】 マルチプロセッサシステムにおいて、より最適なプロセッサ周波数管理システムを提供する。

【解決手段】 マルチプロセッサシステム内のプロセッサの周波数管理の一実施形態では、第1の性能レベルを要求する第1のプロセッサは、特定の周波数で動作して、このシャーシの熱・電力バジェットの一部を消費する。第2の性能レベルを要求する第2のプロセッサは、第2の特定の周波数で動作して、この熱・電力バジェットの一部を消費する。マルチプロセッサのコンピュータ・システム内の電力・熱総合バジェットは、異なる動作周波数のもとで維持される。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

マルチプロセッサ・コンピュータシステム内のプロセッサの動作周波数を管理する方法であって、

第1のプロセッサを第1の周波数で動作させるステップであって、該第1のプロセッサは第1の性能レベルを要求し、該第1の性能レベルに基づいて前記マルチプロセッサ・コンピュータシステム内で、熱・電力バジェットの第1の部分を消費する、ステップと、

第2のプロセッサを第2の周波数で動作させるステップであって、該第2のプロセッサは第2の性能レベルを要求し、該第2の性能レベルに基づいて前記マルチプロセッサ・コンピュータシステム内で、前記熱・電力バジェットの第2の部分を消費する、ステップと

10

前記マルチプロセッサ・コンピュータシステム内の前記第1の周波数と前記第2の周波数の少なくとも1つを調整することで、該マルチプロセッサ・コンピュータシステム内の前記熱・電力バジェットを維持し、該マルチプロセッサ・コンピュータシステムにおける性能を最適化するステップと、

を有する方法。

**【請求項 2】**

前記第1のプロセッサは、前記第2のプロセッサよりも高い性能レベルを要求するプロセスを実行し、また、前記第1のプロセッサは、前記第2のプロセッサよりも高い周波数で動作する請求項1記載の方法。

20

**【請求項 3】**

前記第2のプロセッサは、前記第1のプロセッサよりも低い性能レベルを要求するプロセスを実行し、また、前記第2のプロセッサは、前記第1のプロセッサよりも低い周波数で動作する請求項1記載の方法。

**【請求項 4】**

前記マルチプロセッサ・コンピュータシステムは、P A - R I S C、D E C A l p h a、M I P S、P o w e r P C、S P A R C、I A - 3 2、I A - 6 4から成るグループから選択されたアーキテクチャに基づいている請求項1記載の方法。

**【請求項 5】**

前記熱・電力バジェットの割当ては、前記マルチプロセッサ・コンピュータシステム内の前記第1のプロセッサと前記第2のプロセッサの少なくとも1つの性能要件の変更により、変更される請求項1記載の方法。

30

**【請求項 6】**

マルチプロセッサ・コンピュータシステムであって、

第1の電力割当てを要求するアプリケーションをホストし、第1の周波数で動作し、該第1の電力割当てに基づいて、前記マルチプロセッサ・コンピュータシステム内で、熱・電力バジェットの第1の部分を消費する第1のプロセッサと、

第2の電力割当てを要求するアプリケーションをホストし、第2の周波数で動作し、該第2の電力割当てに基づいて、前記マルチプロセッサ・コンピュータシステム内で、前記熱・電力バジェットの第2の部分を消費する第2のプロセッサと、  
を有し、

40

前記マルチプロセッサ・コンピュータシステム内の前記熱・電力総合バジェットを維持し、また、前記第1の周波数と前記第2の周波数の少なくとも1つを調整することで該マルチプロセッサ・コンピュータシステムにおける性能を最適化する、マルチプロセッサ・コンピュータシステム。

**【請求項 7】**

前記マルチプロセッサ・コンピュータシステムは、P A - R I S C、D E C A l p h a、M I P S、P o w e r P C、S P A R C、I A - 3 2、I A - 6 4から成るグループから選択されたアーキテクチャに基づいている請求項6記載のマルチプロセッサ・コンピュータシステム。

50

**【請求項 8】**

マルチプロセッサ・コンピュータ内のプロセッサの周波数管理のためのシステムであつて、

第1の電力割当てを要求するアプリケーションをホストし、第1の周波数で動作し、該第1の電力割当てに基づいて、前記マルチプロセッサ・コンピュータ内で、熱・電力バジエットの第1の部分を消費するマスタ・プロセッサと、

第2の電力割当てを要求するアプリケーションをホストし、第2の周波数で動作し、該第2の電力割当てに基づいて、前記マルチプロセッサ・コンピュータ内で、前記熱・電力バジエットの第2の部分を消費するスレーブ・プロセッサと、

前記マスタ・プロセッサと前記スレーブ・プロセッサの少なくとも1つの周波数レベルを維持する管理手段と、  
を有するシステム。

**【請求項 9】**

前記マルチプロセッサ・コンピュータは、P A - R I S C 、 D E C A l p h a 、 M I P S 、 P o w e r P C 、 S P A R C 、 I A - 3 2 、 I A - 6 4 から成るグループから選択されたアーキテクチャに基づいている請求項8記載のシステム。

**【請求項 10】**

熱・電力バジエットの割当ては、前記マルチプロセッサ・コンピュータ内の前記マスタ・プロセッサおよび前記スレーブ・プロセッサに対する性能要件の変更により、変更される請求項8記載のシステム。

**【発明の詳細な説明】****【技術分野】****【0 0 0 1】**

本発明は、マルチプロセッサシステムのプロセッサ管理に関する。

**【背景技術】****【0 0 0 2】****(関連出願)**

本願は、A n d r e w H . B A R R 氏らによってなされた「性能要件に基づく、ブレード型アーキテクチャにおけるブレードの周波数管理のためのシステム、方法、装置」と称する米国特許出願第10/216, 437号、A n d r e w H . B A R R 氏らによってなされた「キャパシティ・オン・デマンドを可能にするコンピュータシステムの周波数管理のためのシステムおよび方法」と称する米国特許出願第10/216, 438号、R i c a r d o E S P I N O Z A - I B A R R A 氏らによってなされた「ブレード型システムにおいて、ロードにより周波数および性能を調整するシステムおよび方法」と称する米国特許出願第10/216, 234号、A n d r e w H . B A R R 氏らによってなされた「性能要件に基づく、ブレード型アーキテクチャにおけるブレードの電圧管理」と称する米国特許出願第10/216, 284号、A n d r e w H . B A R R 氏らによってなされた「プロセッサ・レベルでの性能最適化と共同する電圧調整」と称する米国特許出願第10/216, 286号、R i c a r d o E S P I N O Z A - I B A R R A 氏らによってなされた「プロセッサまたはブレードの動作周波数を管理するシステムおよび方法」と称する米国特許出願第10/216, 285号、R i c a r d o E S P I N O Z A - I B A R R A 氏らによってなされた「ブレード型システム内のブレードの動作周波数を管理するシステムおよび方法」と称する米国特許出願第10/216, 229号、A n d r e w H . B A R R 氏らによってなされた「ローディングに基づく、ブレード型システム内のプロセッサの電圧管理」と称する米国特許出願第10/216, 233号、A n d r e w H . B A R R 氏らによってなされた「性能と電力消費を最適化するプロセッサの電圧管理のためのシステムおよび方法」と称する米国特許出願第10/216, 232号、および、A n d r e w H . B A R R 氏らによってなされた「メモリ・サブシステムの管理」と称する米国特許出願第10/216, 435号に關係がある。以上の米国特許出願はすべて、2002年8月12に出願されたものである。

**【0003】**

マルチプロセッサ（MP）アーキテクチャでは、同一システム内に複数のプロセッサがある。特定の組の電力と熱の要件が、このシステムに対応付けられる。公知の電力制限戦略には、CPU機能ユニット、例えば、浮動小数点ユニットまたはオンダイ（on-die）・キャッシュを停止するか、あるいは、ハードディスク・ドライブにおいて、速度と引き換えに、電力消費を削減する方法がある。厳しい電力バジェットおよび熱の制限は、マルチプロセッサ・システム内のプロセッサが動作する最高周波数も制限する場合がある。したがって、これらのプロセッサの最適な性能およびキャパシティが制限される。具体的に言えば、これらの要件は、これらのプロセッサが消費できる電力量に制限を加える。

**【0004】**

MPシステムは、多くの利点を提供するとはいえるが、MPシステムを利用しているときには、いくつかの技術課題が発生する。これらの課題の中には、このシステムにおいて充分に熱を放散させるように、このシステムを設計し、動作させるという課題がある。放熱の課題に取り組むために、MPシステムは、基本的な電力・熱包絡線の範囲内に設計される。例えば、MPシステムをホストするシャーシに、プロセッサの冷却に利用できるエアフローが限られた量しかないとき（すなわち、このシステムが、限られた量の熱だけしか放散させることができないとき）には、このシャーシは、限られた量の電力消費、および、それに伴うプロセッサの制限された性能向けに設計される。

**【0005】**

以前のソリューションには、シャーシの電力・熱冷却総合バジェットを満たすために、最適よりも低い性能レベルで、これらのプロセッサを動作させること、ファンおよび特別の制御回路を追加すること、このシステム内のI/Oカードまたは他の有用な機構の数を制限すること、および、このシステム内の他の機構に使える電力バジェットを削減することがあった。

**【発明の開示】****【発明が解決しようとする課題】****【0006】**

本発明の目的は、マルチプロセッサシステムにおいて、より最適なプロセッサ周波数管理システムを提供することである。

**【課題を解決するための手段】****【0007】**

マルチプロセッサ（MP）システム内のプロセッサの周波数管理の一実施形態では、第1の性能レベルを要求する第1のプロセッサは、特定の周波数で動作して、このシャーシの熱・電力バジェットの一部を消費する。第2の性能レベルを要求する第2のプロセッサは、第2の特定の周波数で動作して、この熱・電力バジェットの一部を消費する。マルチプロセッサのコンピュータ・システム内の電力・熱総合バジェットは、異なる動作周波数のもとで維持される。

**【0008】**

マルチプロセッサ・システム内のプロセッサの周波数管理の別の実施形態では、第1の性能レベルを要求するマスタ・プロセッサは、特定の周波数で動作して、この熱・電力バジェットの一部を消費する。第2の性能レベルを要求するスレーブ・プロセッサは、第2の特定の周波数で動作して、この熱・電力バジェットの一部を消費する。さらに、管理手段を用いて、マスタ・プロセッサおよび／またはスレーブ・プロセッサの周波数レベルを維持する。

**【発明を実施するための最良の形態】****【0009】**

次に、同じ要素に同じ番号を付した以下の図を参照して、マルチプロセッサ（MP）システム内のプロセッサの周波数管理の好ましい実施形態を詳しく説明する。添付図面のうちの図1を参照すると、ブレード型アーキテクチャ・システムの基本モジュラー構成要素の一実施形態を描いたブロック図（全体が参照数字100で示される）が図1に示されて

いる。ブレード型アーキテクチャは、この開示が有益となりうる多くの異なるタイプのコンピュータ・アーキテクチャの一例である。本明細書に述べられる新技術は、様々なMPサーバーまたはMPコンピュータに適用できることが理解されよう。

#### 【0010】

管理ブレード110は、このシャーシの機能を監視して、インストールされたあらゆるサーバーのコンソールとの单一インターフェースを提供する。図1に示されるように、サーバー・ブレード120は、管理ブレード110と通信している。次に、サーバー・ブレード120が、特定の機能を果たす他のブレードと通信している。例えば、図1に見られるように、サーバー・ブレード120は、ファイバ・チャネル・ブレード130およびネットワーク・ブレード140と通信している。ブレード型アーキテクチャ・システム内の様々なブレードは、プロセッサ・ブレード、サーバー・ブレード、ネットワーク・ブレード、ストレージ・ブレード、またはストレージ相互接続ブレードなどである場合もあると理解されよう。

#### 【0011】

プロセッサ・レベルでの性能最適化は、マルチプロセッサ(MP)装置内の個々のプロセッサを、それらの特定の性能レベルおよびニーズにより設定できるようにすることで、リソースのインテリジェント利用を可能にする。性能を上げたり、または下げたりするためのプロセッサの要件を利用して、それぞれのプロセッサに、高くした周波数または低くした周波数で動作させ、したがって、それぞれのプロセッサは、シャーシの熱・電力バジエットのうち、消費する量が多くなるか、または少なくなる。

#### 【0012】

さらに低い性能レベルを要求するバックグラウント・プロセスを実行するプロセッサ、例えば、スレーブ・プロセッサは、さらに低い周波数で動作し、したがって、シャーシの熱・電力バジエットのうち、消費する量が少なくなる。さらに高い性能レベルを要求するプロセスを実行するプロセッサは、さらに高い周波数で動作し、したがって、シャーシの熱・電力バジエットのうち、消費する量が多くなる。いずれのシナリオにおいても、システムの熱・電力総合要件は、やはり、さらに最適な総合プロセッサ性能を用いて満たされる。PAアーキテクチャでは、そのプロセッサ・コア周波数は、バス動作周波数と非同期である。したがって、プロセッサ・コア周波数に変更を加えることができ、しかも、その影響はほとんどない。当業者であれば、このような原理は、DEC Alpha、MIPS、PowerPC、SPARC、IA-32、IA-64のアーキテクチャ、および他のMPアーキテクチャからの他のプロセッサにも応用できることがすぐ理解されよう。

#### 【0013】

図2～図7は、MPシステム内のプロセッサの周波数管理のための様々な方法を示している。いくつかのアーキテクチャでは、CPUのプロセッサ・コア周波数は、バス動作周波数と非同期であり、それゆえ、システムバス周波数とは無関係に、プロセッサ・コア周波数に変更を加えることができる。現在、ヒューレット・パッカード社は、PARISCシステムのプロセッサ・クロックを発生させる周波数シンセサイザとして、SPHYR-T ASICを使用している。しかしながら、当業者であれば、他の周波数シンセサイザも利用できることがすぐ理解されよう。プロセッサ・アーキテクチャの他の実施形態、例えばIA-32やIA-64では、CPUのプロセッサ・コア周波数は、必ずしも、バス動作周波数と非同期であるとは限らない。IAのプロセッサでは、プロセッサ・コア周波数は、バス動作周波数の設定可能な倍数で動作する。このクロック・チップの出力を変更し、したがって、それに応じてプロセッサ周波数をセットすれば、個々のプロセッサの動作周波数の管理を、双方のアーキテクチャに適用することができる。

#### 【0014】

一般に、プロセッサのクロックを発生させるのに用いられる周波数シンセサイザ・チップは、パラレル又はシリアルの構成ビットを持ち、それらの構成ビットにより、出力クロック(合成周波数)に対する入力クロックの比率を選択することができる。これにより、リブートのときに、プロセッサを異なる周波数で動作させることができる。この周波数シ

ンセサイザは、通常、コア・クリスタルからの入力周波数を備えている。シリアル・ピンまたはパラレル・ピンの制御を通じて、この周波数シンセサイザは、プロセッサに送られる出力周波数比率を提供する。

#### 【0015】

当業者であれば、複数、すなわち2個（2方向システム用）、またはN個（N方向システム用）の周波数シンセサイザを利用して、プロセッサのクロックを発生させる場合に、それらのプロセッサは、プロセッサでサポートされている周波数の範囲内の異なる周波数で動作できることも理解されよう。図2～図7は、異なる周波数シンセサイザの比率ピン（ratio pin）をどのように制御できるのか、異なる手法を示している。

#### 【0016】

次に、添付図面のうちの図2を参照すると、手動設定装置210を用いてMPシステム内の個々のプロセッサの動作周波数を管理する一手法を描いたブロック図（全体が参照数字200で示される）が示されている。図2に見られるように、このシステムの容易に手が届く部分に、手動設定装置210が付け加えられている。したがって、所定の性能要件に基づいて、このユニットのリブートのときに、プロセッサの周波数を操作者にセットさせる。通常の当業者であれば、所望の機能を果たすことのできる多くの共通手動設定装置、例えばディップ・スイッチ、ピン・ヘッダの上に設けられるジャンパ、回転設定スイッチ、はんだプリッジなどがあることが理解されよう。入力周波数230と、手動設定装置210からの信号220を使用して、クロック・ジェネレータまたは周波数シンセサイザ240で出力周波数250を発生させ、その出力周波数250を、プロセッサ260で使用する。手動設定装置を用いてプロセッサの周波数を変更する前述の方法、および、本明細書中の下記の他の方法は、様々なタイプのMPアーキテクチャに適用できるものと理解されよう。

#### 【0017】

次に、添付図面のうちの図3を参照すると、MPシステム内の個々のプロセッサの動作周波数を管理する他の方法を描いたブロック図（全体が参照数字300で示される）が示されている。方法300は、抵抗器310を用いて、個々のプロセッサの動作周波数を管理している。当技術分野で知られているように、抵抗器310は、設定用抵抗器であることもある。図3に見られるように、このシステムの容易に手が届く部分に、抵抗器310が付け加えられている。したがって、所定の性能要件に基づいて、このユニットのリブートのときに、プロセッサの周波数を操作者にセットさせる。入力周波数330と、抵抗器310からの信号320を使用して、クロック・ジェネレータ340で出力周波数350を発生させ、その出力周波数350を、プロセッサ360で使用する。

#### 【0018】

次に、添付図面のうちの図4を参照すると、マイクロコントローラまたはマイクロプロセッサ420を用いてMPシステム内のプロセッサの動作周波数を管理する方法を描いたブロック図（全体が参照数字400で示される）が示されている。マイクロコントローラまたはマイクロプロセッサ420は、このシステムのユーザと対話して、それぞれのプロセッサが動作すべき特定の周波数を求めるために用いられる。図4に見られるように、マイクロコントローラまたはマイクロプロセッサ420は、GSP（ガーディアン・サービス・プロセッサ）または他のコントローラからのI<sup>2</sup>C（Inter-IC）バス410で信号を受け取る。当技術分野で知られているように、I<sup>2</sup>Cバスは、集積回路間に通信リンクを提供する双方向の2線式シリアル・バスである。さらに、通常の当業者であれば、他の制御バスであっても、同一機能を果たして、本明細書に述べられるI<sup>2</sup>Cバスに置き換えることも、すぐ理解されよう。マイクロコントローラまたはマイクロプロセッサ420は、ユーザにより指定された特定の周波数に基づいて、パラレルまたはシリアルの制御信号430を出力する。クロック・ジェネレータ450は、入力周波数440と、パラレルまたはシリアルの制御信号430を利用して、指定したプロセッサ470で用いられる出力周波数460を発生させる。マイクロコントローラまたはマイクロプロセッサ420を使えば、ユーザは、上述のレジスタと手動設定装置の方法よりも透過的なやり方

で、周波数シンセサイザを制御することができる。すなわち、ユーザは、コンフィグレーション・ビットのセッティングが、どのように、その出力に影響を及ぼすのか、必ずしも知る必要はない。

#### 【0019】

次に、添付図面のうちの図5を参照すると、FPGA (field-programmable gate array, 書替可能ゲートアレイ) またはPLD (programmable logic device, プログラム可能論理素子) 520を用いてMPシステム内のプロセッサの動作周波数を管理する方法を描いたブロック図（全体が参照数字500で示される）が示されている。当業者で知られているように、FPGAは、製造後に、現場でプログラムできるチップである。FPGAまたはPLD520は、I<sup>2</sup>Cバス510を通じて、さらに高いレベルの装置（例えば、GSP）からコマンドを受け取って、その周波数シンセサイザ（または、それぞれのシンセサイザ）に対するコンフィグレーション・ビットを制御するために用いられる。図5に見られるように、クロック・ジェネレータ550は、入力周波数540と、FPGA/PLD530からの制御信号を利用して、プロセッサ570で用いられる出力周波数560を発生させる。図4で説明されるマイクロコントローラ/マイクロプロセッサ420の利用と同様に、FPGA/PLD520を用いれば、ユーザは、さらに透過的なやり方で、周波数シンセサイザを制御することができる。すなわち、ユーザは、コンフィグレーション・ビットのセッティングが、どのように、その出力に影響を及ぼすのか、必ずしも知る必要はない。

#### 【0020】

次に、添付図面のうちの図6を参照すると、I/Oエキスパンダ・チップ（I/OX）620を用いてプロセッサの動作周波数を管理する方法を描いたブロック図（全体が参照数字600で示される）が示されている。当技術分野で知られているように、I<sup>2</sup>CベースのI/OX620は、それぞれの周波数シンセサイザを透過的に制御するのに使用できる安価で、かつ単純なソリューションである。I/OXチップ620は、I<sup>2</sup>Cコマンドを通じてI/OXに書き込めば、特定の状態に強制できるI/Oポートを持っている。I/OXチップは、通常、複数のI/Oポートを持っているから、1つのI/OXを使用すれば、複数の周波数シンセサイザを個々に制御できるものと理解されよう。

#### 【0021】

I/OXチップはI<sup>2</sup>Cベースのものであるから、I/OXチップは、I<sup>2</sup>Cインターフェースをサポートする装置であれば、どんなものによっても制御できる。ヒューレット・パッカード社のPA-RISCブレードでは、このような論理装置は、このPAブレードのサービス・プロセッサすなわちGSPであろう。GSPのユーザ・フレンドリなインターフェースを使用すれば、どんな低レベルの情報（例えば、ビット・セッティング）も必要とせずに、プロセッサの周波数を顧客に、透過的に制御させることができる。GSPからの入力は、図6に、参照数字610で示されている。ユーザは、GSPにて、コマンドを入力して、プロセッサの周波数を高める。次に、GSPは、プロセッサが所望のレベルで動作するように、所要のビットを受け渡して周波数発生器の比率を変更する。ビット・ストリーム610が、I/OX620で受け取られる。次に、クロック・ジェネレータ650は、入力周波数640と制御信号630を使用して、プロセッサ670で用いられる出力周波数660を発生させる。

#### 【0022】

次に、添付図面のうちの図7を参照すると、IAベースのアーキテクチャ・システム内の個々のプロセッサの動作周波数を管理する方法を描いたブロック図（全体が参照数字700で示される）が示されている。上に示されるように、PAベースのアーキテクチャは、プロセッサに取り入れられる2つの周波数入力、すなわちプロセッサ・コア入力とバス入力を持っている。IAベースのアーキテクチャでは、シリアルまたはパラレルの制御信号は調整されない。IAベースのアーキテクチャは、バス周波数の倍数を発生させるように、プロセッサ・コア周波数をプログラムする。このようにプロセッサ・コア周波数をプログラムする作業は、多目的バス・ライン730を用いて、当初のバス初期設定の間に行

われる。これは、通常、プロセッサ・バス（C E C）720上にある主コア・チップセットによって制御される。一般に、ユーザ・インターフェース、例えばG S P 710を使用して、当初のシステムおよびバスの初期設定の間に、適切な制御信号730を発生させるようにC E C 720をプログラムする。特定のプロセッサ760におけるアプリケーションに応じて、さらに大きい比率をバス周波数に掛けて、さらに高い周波数、性能、電力を発生させるか、あるいは、さらに小さい比率をバス周波数に掛けて、さらに低い周波数、性能、電力を発生させる。

#### 【0023】

次に、添付図面のうちの図8を参照すると、同一周波数で動作するM Pシステム内の一連のプロセッサを描いたブロック図（全体が参照数字800で示される）が示されている。個々のプロセッサの陰影は、個々のプロセッサが、それぞれ同じ周波数レベルで動作していることを示している。同一周波数でプロセッサを動作させることができ、現行のM Pアーキテクチャ・システムでは代表的なものである。さらに、この陰影は、このシステム全体に割り当てられた最大電力未満にとどまるように、それぞれのプロセッサが、最高レベルよりも低いレベルで動作していることを示している。論じられたように、M Pサーバー・システムは、基本的な電力・熱包絡線によって制限される。これは、生じた熱や、このシステムにおける限られた寸法によるものである。このシステムが、所与の電力量を消費するときには、このシステムは、通常、プロセッサを冷却するのに利用できるエアフローの量が制限される。その結果、このような電力制限は、プロセッサが動作できる周波数を制限し、したがって、その性能を制限する。それゆえ、プロセッサは、最適な性能およびキャパシティで動作できる能力が制限される。なぜなら、これらのプロセッサは、同一周波数、すなわち、それらの最高レベルよりも低い周波数で動作するように設定されているからである。

#### 【0024】

次に、添付図面のうちの図9を参照すると、異なる周波数で動作するM Pシステム内のプロセッサを描いたブロック図（全体が参照数字900で示される）が示されている。図9に示されるプロセッサ周波数の新たな設定の根拠は、M P装置内の個々のプロセッサを、それらの特定の性能レベルおよびニーズにより設定できるようにすることで、リソースのインテリジェント利用に置かれている。この図の陰影は、プロセッサが動作している周波数のレベルを示している。例えば、陰影がさらに多いプロセッサは、さらに高い周波数で動作している。

#### 【0025】

性能を上げたり、または下げたりするためのプロセス要件を利用して、それぞれのプロセッサを、高くした周波数または低くした周波数で動作させ、したがって、それぞれのプロセッサは、シャーシの熱・電力バジェットのうち、消費する量が多くなるか、または少なくなる。バックグラウント・プロセスを実行するスレーブ・プロセッサは、さらに低い性能レベルを要求して、さらに低い周波数で動作する。したがって、これらのスレーブ・プロセッサは、シャーシの熱・電力バジェットのうち、消費する部分が少なくなる。さらに高い性能レベルを要求するプロセスを実行するマスター・プロセッサすなわちモナーク・プロセッサ（monarch processor）は、さらに高い周波数で動作する。したがって、マスター・プロセッサは、シャーシの熱・電力バジェットのうち、消費する部分が多くなる。さらに、これらのプロセッサに対する相対性能要件が変わると、周波数を変更でき、したがって、性能、および熱・電力バジェットの割当てが変えられる。やはり、個々のプロセッサを異なる周波数で動作させることにより、このシステムの熱・電力総合要件が満たされて、さらに最適な総合プロセッサ性能が得られる。

#### 【0026】

本明細書に開示される原理は、共通シャーシを共有するプロセッサから成るシステムに、あるいは、複数のシャーシにまたがるアーキテクチャ・システムに応用できるものと理解されよう。すなわち、このような原理は、物理パーティションまたは論理パーティションで分割されたシステムに応用できる。例えば、物理的に、システムは、それぞれ8基の

プロセッサを有する3つのシャーシを含むことがある。論理的に、同一システムが、5人の異なるカスタマに対して、5つの異なるwebサーバーに区分されることもある。シャーシ内の電力制約は、通常、このシステムの物理パーティションに関係する。複数のシャーシ内に位置づけられるカスタマまたはアプリケーションに課せられる電力制約は、通常、論理パーティションに関係する。通常の当業者であれば、上述の技術革新は、物理的にも論理的にも区分されたアーキテクチャに適用できることがすぐ理解されよう。

#### 【0027】

MPシステム内のプロセッサの周波数管理が、模範的な実施形態に関連して述べられてきたが、当業者であれば、以上の教示に照らして多くの変更が可能であり、また、本願が、そのどんな変形例もカバーするようにもくろまれていることが理解されよう。

10

#### 【0028】

例えば、この開示されたシステムおよび方法は、上記の実施形態の一部において、概ね、ブレード型アーキテクチャ・システムに適用されてきた。他のコンピュータ・アーキテクチャも、同じように用いられることがある。したがって、この開示全体を通じて、図示され、参照されるMPアーキテクチャは、特記しない限り、所望の機能を果たすのに適した一切のアーキテクチャを表わすことになっている。同様に、様々な動作を行ういくつかのI<sup>2</sup>C装置もいくつか開示されている。この特定のI<sup>2</sup>C装置は、この開示を制限するつもりはない。したがって、この開示を、任意の特定の形態のI<sup>2</sup>C装置、または特定のアーキテクチャに限定することは、出願者の意図ではない。

#### 【0029】

この開示全体を通じて、さらに他の例証があり、本明細書にはつきりと特定されてはいないが、それでも、クレーム記載の機能を果たすことのできる構造、材料、または動作の利用を、この開示の範囲から除外することは、出願者の意図ではない。

20

#### 【図面の簡単な説明】

#### 【0030】

【図1】ブレード型アーキテクチャ・システムの基本モジュラー構成要素の一実施形態を描いたブロック図。

【図2】手動設定装置を用いて、個々のプロセッサの動作周波数を管理する一手法を描いたブロック図。

30

【図3】抵抗器を用いて、個々のプロセッサの動作周波数を管理する他の方法を描いたブロック図。

【図4】マイクロコントローラまたはマイクロプロセッサを用いて、個々のプロセッサの動作周波数を管理する他の方法を描いたブロック図。

【図5】FPGAまたはPLDを用いて、個々のプロセッサの動作周波数を管理する他の方法を描いたブロック図。

【図6】I/OXを用いて、個々のプロセッサの動作周波数を管理する他の方法を描いたブロック図。

【図7】IAベースのアーキテクチャ・システム内の個々のプロセッサの動作周波数を管理する方法を描いたブロック図。

【図8】同一周波数レベルで動作するMPシステム内の一連のプロセッサを描いたブロック図。

40

【図9】異なる周波数レベルで動作するMPシステム内の一連のプロセッサを描いたブロック図。

#### 【符号の説明】

#### 【0031】

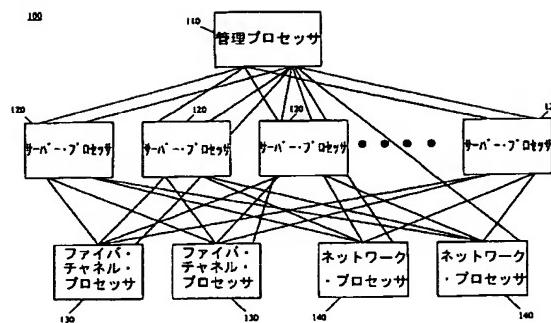
100 マルチプロセッサ・コンピュータシステム

110 マスター・プロセッサ

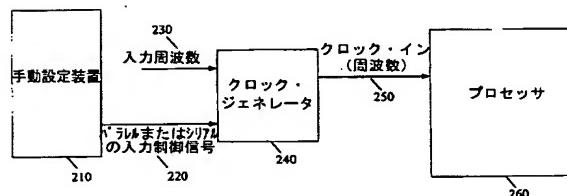
120 第1のプロセッサ、スレーブ・プロセッサ

130 第2のプロセッサ

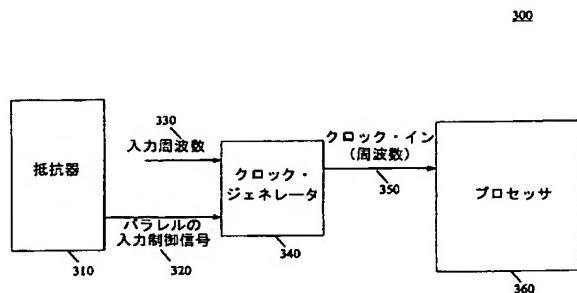
【図 1】



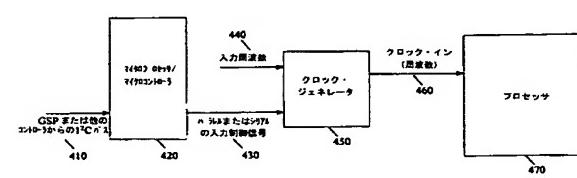
【図 2】



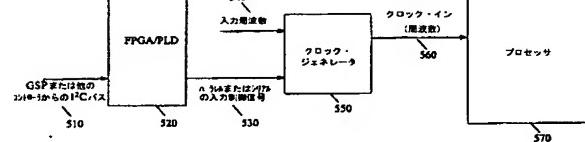
【図 3】



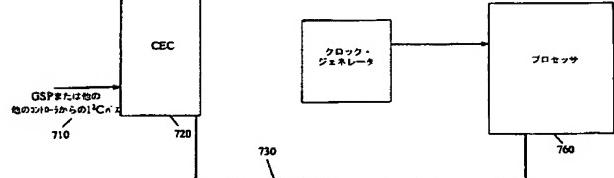
【図 4】



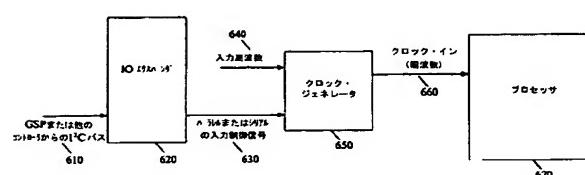
【図 5】



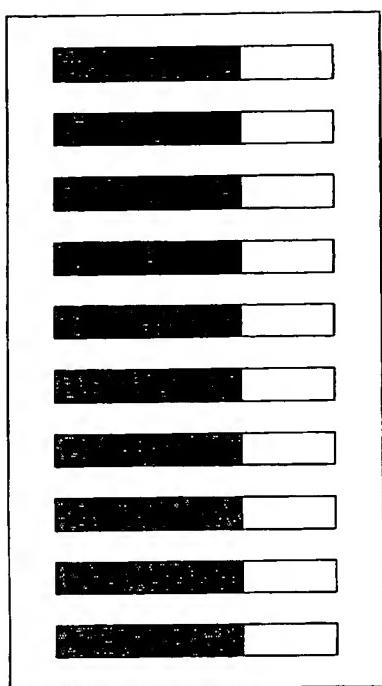
【図 7】



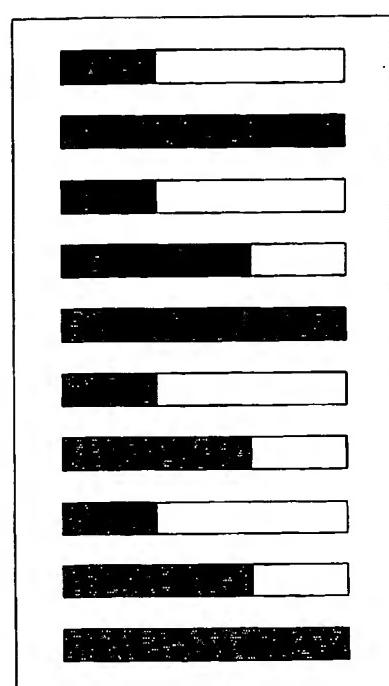
【図 6】



【図8】

800

【図9】

900

---

フロントページの続き

(72)発明者 リカルド・エスピノサーイバーラ  
アメリカ合衆国95608カリフォルニア州カーマイケル、フェアウッド・ウェイ 4004  
(72)発明者 アンドリュー・エイチ・バー<sup>1</sup>  
アメリカ合衆国95678カリフォルニア州ローズヴィル、モンゴメリ・ストリート 117  
Fターム(参考) 5B045 CC01 KK01 KK02  
5B079 AA06 BA01 BC01